香港天文台
HONG KONG OBSERVATORY

Reprint 1342

Probabilistic forecasts for annual mean temperature reaching the highest top 10 and top 5 positions

CHAN Wing-hang, TONG Hang-wai, LEE Sai-ming

The 32st Guangdong - Hong Kong - Macao Seminar
on Meteorological Science and Technology
and
The 23nd Guangdong - Hong Kong - Macao Meeting
on Cooperation in Meteorological Operations

(Macau 8-10 January 2018)

年平均氣溫列入最高前十名和前五名的概率預報

陳永鏗　唐恒偉　　李細明
香港天文台

摘要

受全球暖化和本地城市化的影響，香港天文台總部的溫度有長期而且顯著的上升趨勢。因此過去 20 年有多次年平均氣溫處於最高前十名甚至前五名之列。作為年初首季對全年氣溫趨勢的一項評估，天文台利用一、二月的實況觀測，結合 ECMWF (三月到九月)及 NCEP(三月到十一月)經校正的預測和餘下月份的氣候平均，發展年平均氣溫達致最高前十名和前五名的概率預報。基於布賴爾技巧評分(Brier Skill Score)而計算的 1997-2016 年驗証結果顯示這個預報方法是有效可行。由概率預報轉換成決定性預報的表現亦令人滿意。

Probabilistic forecasts for annual mean temperature reaching the highest top 10 and top 5 positions

CHAN Wing-hang　TONG Hang-wai　LEE Sai-ming
Hong Kong Observatory

Abstract

Under the influence of global warming and local urbanization, the temperature at the Hong Kong Observatory (HKO) headquarters reveals a significant long-term rising trend. The observed annual mean temperatures are among the highest top 10 or even top 5 positions for many times over the past two decades. As an assessment of the annual temperature trend during the first quarter in the early part of the year, HKO made use of the actual temperature observations in January and February and combined them with the calibrated forecasts from ECMWF for March to September and from NCEP for March to November and also the climatology for the remaining month(s) to develop probabilistic forecasts of annual mean temperature reaching the highest top 10 and top 5 positions. Verification for the forecasts in 1997-2016 using Brier Skill Score shows that the methodology adopted is feasible and applicable. Conversion from probabilistic forecasts to deterministic event forecasts also produces satisfactory performance.

1. Introduction

Under the influence of global warming and local urbanization, the temperature at the Hong Kong Observatory (HKO) headquarters exhibits a significant warming trend. Many of the top 10 and top 5 hottest years were registered in the past two decades (Table 1). There is an increasing media and public interest in the occurrence of record breaking temperature. HKO made use of the actual temperature observations in January and February and combined them with the calibrated forecasts from ECMWF for March to September and from NCEP for March to November and also the climatology of the remaining month(s) to develop probabilistic forecasts of annual mean temperature reaching the highest top 10 and top 5 positions (TOP10 and TOP5[1]). Section 2 describes the methodology while the verification results are presented in Section 3, followed by a conclusion in Section 4.

2. Methodology

2.1 Deterministic annual mean temperature forecast

Since the annual outlook is usually issued in late March, January and February temperatures will be available by then. Calibrated ECMWF and NCEP ensemble hindcast/forecast data initialized in late February or early March for the period 1981/1982-2016 (1981 for ECMWF, 1982 for NCEP) were utilized in this study (Table 2). The validation time of ECMWF and NCEP hindcast/forecast data covers March to September and March to November respectively. Combining the observations, averaged calibrated model data and the climate normal of the remaining months together will form the deterministic annual mean temperature forecast for the year (Figure 1).

2.2 Calibration of model data

It is known that systematic bias exists in ensemble forecast and hence calibration is needed [1]. In this study, three calibration methods were applied to each ensemble member of model data month by month, namely quantile-quantile mapping (QQM), standardized anomaly mapping (SAM) and linear regression. The calibrated members were then averaged for each month and each model. In QQM, the relative position of the predictor (direct temperature output given by models) in a training data set is first determined. The value of the predictand (temperature at HKO) with the same position in the predictand's training data set was then taken as the forecast. In SAM, the standardized anomaly of the predictor was taken as the standardized anomaly forecast of the predictand. The basic training period was 1981/1982-1996 and the verification period was 1997-2016. As the verification year progressed from 1997 to 2016, the training period would extend year by year accordingly.

[1] TOP10 (TOP5) denotes the event of annual mean temperature at HKO reaching the highest top 10 (top 5) positions.

2.3 Probabilistic forecast of TOP10 and TOP5

In order to generate a probabilistic forecast, a normal distribution centred at the annual mean temperature forecast was assumed. The standard deviation of the normal distribution was estimated by the root-mean-squared error (RMSE). An initial estimate was given by the cross-validation results for the basic training data (1981/1982-1996). The estimate would then be updated year by year accordingly as the verification year progressed from 1997 to 2016. With the parameters of the temperature distribution specified, we can work out the probability of TOP10 and TOP5.

To facilitate skill evaluation, a reference forecast was constructed from the observed January and February temperatures of the verification year and the climatological data of March-December temperature. The March-December temperature was assumed to follow a normal distribution with parameters estimated from the climatological data.

2.4 Deterministic event forecast of TOP10 and TOP5

There may be occasions that a deterministic event forecast is preferred by members of the public. The probabilistic forecast was converted to a deterministic event forecast according to the following algorithm:

(a) If the probability was 50% or above, forecast occurrence of TOP10/TOP5.
(b) Otherwise, forecast no occurrence of TOP10/TOP5.

Metrics including percentage correct, hit rate, false alarm rate and Critical Success Index (CSI, [2]) were used to evaluate the performance of the deterministic event forecast.

3. Verification Results

Brier Skill Score (BSS, [3]) was evoked to gauge the performance of the probabilistic forecast:

$$BSS = 1 - \frac{BS}{BS_{ref}}$$

where BS and BS_{ref} are the Brier Score achieved by the probabilistic forecast and the reference forecast respectively. Brier Score is a quadratic scoring measure for probabilistic forecast [3]. The interpretation of BSS is as follows:

BSS = 1 :	The forecast system is perfect
0 < BSS < 1:	The forecast system performs better than the reference forecast
BSS = 0 :	The forecast system is as good as the reference forecast
BSS < 0 :	The forecast system is inferior to the reference forecast

Table 3 shows that all probabilistic forecasts are skillful compared to the reference forecast. For TOP10, the average of two models achieved the highest BSS of 0.45. For TOP5, ECMWF achieved the highest BSS, followed by the average of two models.

Figure 2 shows the probabilities of TOP10 given by ECMWF, NCEP and the reference forecast. The black solid line is the observed annual mean temperature at HKO. The red solid line is the 10^{th} highest annual mean temperature for data records ending at the preceding year. When the black line touched or was above the red line, TOP10 occurred. As shown by the figure, the probabilities given by the models reflected the likelihood of the event in most of the years. Similar observations could be made in Figure 3 for TOP5 in 1997-2006, but results were less successful in recent years.

Table 4 shows the percentage correct, hit rate, false alarm rate and CSI of deterministic event forecasts of TOP10 given by ECMWF, NCEP and the average of two models. The percentage of correct forecasts exceeded 70% in all cases. Hit rates higher than 70% were achieved with false alarm rates below 20% and corresponding CSIs above 0.65.

Table 5 shows the percentage correct, hit rate, false alarm rate and CSI of deterministic event forecasts of TOP5 given by ECMWF, NCEP and the average of the two models. The percentage of correct forecasts was at least 70% in all cases, but the hit rates dropped to 50-60%. The false alarm rates remained low except for NCEP. The CSIs were at least 0.45 in all cases.

4. Conclusion

A probabilistic forecast of TOP10/TOP5 was developed using actual temperature observations, calibrated model data and climatological data, and evaluated to be skillful. The performance of the deterministic event forecast converted from the probabilistic forecast was satisfactory as well. Taking all the performance evaluation results into account, the average of the two models is considered a suitable candidate for generating the probabilistic forecast and deterministic event forecast of TOP10.

Reference

[1] Cui, B., Z. Toth, Y. Zhu and D. Hou, 2012: Bias correction for global ensemble forecast. *Weather and Forecasting*, **27**, 396-410

[2] Mason, I. B., 2003: *Probability and ensemble forecast. Forecast Verification: A Practitioner's Guide in Atmospheric Science*, I. T. Jolliffe and D. B. Stephenson. Eds., John Wiley and Sons, 37-76

[3] Toth, Z., O. Talagrand, G. Candille and Y. Zhu, 2003: *Probability and ensemble forecast. Forecast Verification: A Practitioner's Guide in Atmospheric Science*, I. T. Jolliffe and D. B. Stephenson. Eds., John Wiley and Sons, 137-163

Table 1 The highest top 10 annual mean temperature (°C) recorded at HKO. Years between 1997 and 2016 are highlighted in red.

No	Rank	Annual Mean Temperature	Year
1	1	24.2	2015
2	2	24.0	1998
3	3	23.9	2002
4	4	23.8	1966
5	4	23.8	1999
6	6	23.7	2007
7	7	23.6	1994
8	7	23.6	2001
9	7	23.6	2003
10	7	23.6	2016

Table 2 Specifications of ECMWF and NCEP hindcast/forecast data

	ECMWF	NCEP (time-lagged)
Initial date	1 March	15, 20, 25 February (hindcast) 15-25 February (forecast)
Forecast months	March to September	March to November
Number of ensemble members	15 (hindcast) 51 (forecast)	12 (hindcast) 44 (forecast)
Hindcast period	1981-2011	1982-2010

Table 3 BSS of probabilistic forecasts of TOP10/TOP5

	ECMWF	NCEP	Average of two models
TOP10	0.43	0.42	0.45
TOP5	0.50	0.37	0.47

Table 4 Percentage correct, hit rate, false alarm rate and CSI of deterministic event forecasts of TOP10

	ECMWF	NCEP	Average of two models
Percentage correct	75%	75%	75%
Hit rate	71%	79%	79%
False alarm rate	9%	15%	15%
CSI	0.67	0.69	0.69

Table 5 Same as Table 4, but for deterministic event forecasts of TOP5

	ECMWF	NCEP	Average of two models
Percentage correct	80%	70%	85%
Hit rate	50%	63%	63%
False alarm rate	0%	38%	0%
CSI	0.50	0.45	0.63

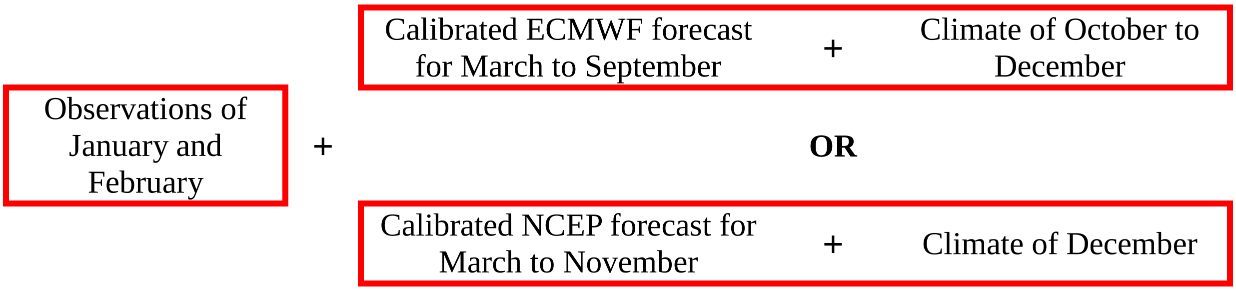

Figure 1 Schematic diagram showing the formulation of deterministic annual mean temperature forecast.

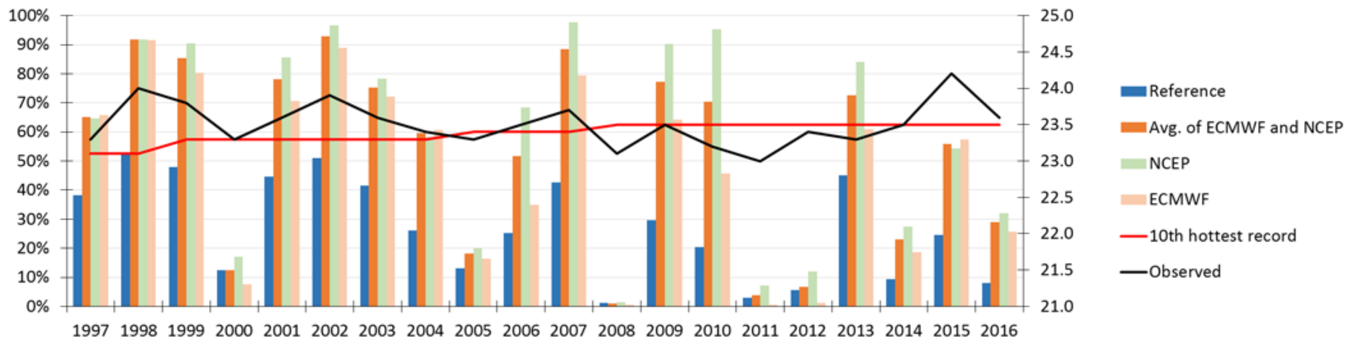

Figure 2 Probabilities of TOP10 given by ECMWF, NCEP, average of the two models and the reference forecast for 1997-2016. The black solid line shows the observed annual mean temperature. The red solid line shows the 10th highest annual mean temperature for data records ending at the preceding year.

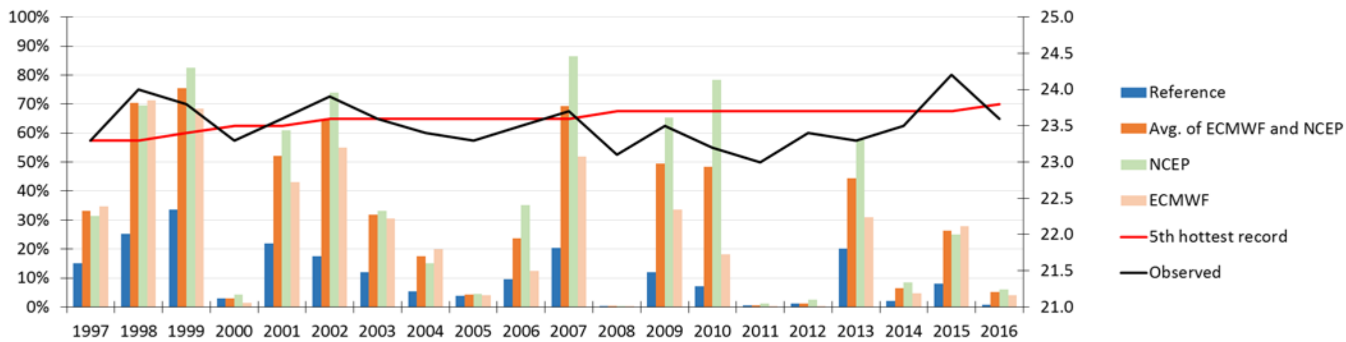

Figure 3 Same as Figure 2, but for probabilities of TOP5. The red solid line shows the 5th highest annual mean temperature for data records ending at the preceding year.